

SEQ SEARCH SUMMARY

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 12, 2004, 13:33:34 ; Search time 67396 Seconds
(without alignments)
17817.359 Million cell updates/sec

Title: US-09-942-025-12
Perfect score: 27705
Sequence: 1 gtggcgccgatggacagaat.....cagaggatctccatgaatag 27705 - nucleotides long
Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 200 — see below

Total number of hits satisfying chosen parameters: 1

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*

* B/c the claims are limited to exact SEQ 12 or a portion that is at least a PKS domain AND b/c in ~~the~~ the art, the smallest PKS domain is greater than 65 amino acids (195 nucleotides), limiting the search to a word size of 200 will retrieve all relevant hits!

28: em_un:*
 29: em_vi:*
 30: em_htg_hum:*
 31: em_htg_inv:*
 32: em_htg_other:*
 33: em_htg_mus:*
 34: em_htg_pln:*
 35: em_htg_rod:*
 36: em_htg_mam:*
 37: em_htg_vrt:*
 38: em_sy:*
 39: em_htgo_hum:*
 40: em_htgo_mus:*
 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	14165	51.1	33529	6	AR166425	AR166425 Sequence

ALIGNMENTS

RESULT 1
 AR166425
 LOCUS AR166425 33529 bp DNA linear PAT 17-OCT-2001
 DEFINITION Sequence 3 from patent US 6280999.
 ACCESSION AR166425
 VERSION AR166425.1 GI:16241741
 KEYWORDS .
 SOURCE Unknown.
 ORGANISM Unknown.
 UNCLASSIFIED
 REFERENCE 1 (bases 1 to 33529)
 AUTHORS Gustafsson,C., Betlach,M.C., Ashley,G., Julien,B. and Ziermann,R.
 TITLE Sorangium polyketide synthases and encoding DNA therefor
 JOURNAL Patent: US 6280999-A 3 28-AUG-2001;
 FEATURES Location/Qualifiers
 source 1. .33529
 /organism="unknown"
 /mol_type="unassigned DNA"

Applicant's CIP parent

ORIGIN

Query Match 51.1%; Score 14165; DB 6; Length 33529;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 15255; Conservative 0; Mismatches 5; Indels 7; Gaps 7;

Qy 12439 GATCCCCAGCAGCGGCTGGTGCTGGAGACGGCGTGGGAGGCATTGGAGCGTGCCGGCGTG 12498
 |||
 Db 1 GATCCCCAGCAGCGGCTGGTGCTGGAGACGGCGTGGGAGGCATTGGAGCGTGCCGGCGTG 60

OM nucleic - nucleic search, using sw model

Run on: June 12, 2004, 13:31:29 ; Search time 5997 Seconds
 (without alignments)
 19625.868 Million cell updates/sec

Title: US-09-942-025-12
 Perfect score: 27705
 Sequence: 1 gtggcgccgatggacagaat.....cagaggatctccatgaatag 27705

Scoring table: OLIGO_NUC
 Gapop 60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 200

Total number of hits satisfying chosen parameters: 7

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : N_Geneseq_29Jan04:*
 1: geneseqn1980s:*
 2: geneseqn1990s:*
 3: geneseqn2000s:*
 4: geneseqn2001as:*
 5: geneseqn2001bs:*
 6: geneseqn2002s:*
 7: geneseqn2003as:*
 8: geneseqn2003bs:*
 9: geneseqn2003cs:*
 10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	25767	93.0	27705	9	ADC26979	Adc26979 Sorangium
2	20582	74.3	67251	9	ADC26995	Adc26995 Sorangium
3	14165	51.1	33529	5	AAS17367-US 6280999	Aas17367 DNA seque
4	752	2.7	20922	9	ADC26981	Adc26981 Sorangium
5	480	1.7	11358	9	ADC26983	Adc26983 Sorangium
6	368	1.3	750	3	AAA58917-US 6090601	Aaa58917 DNA encod
7	220	0.8	787	3	AAA58916- " "	Aaa58916 DNA encod

App's US 2003/054547
no mismatches in US PG Pubs → must be errors in Geneseq

OM nucleic - nucleic search, using sw model

Run on: June 12, 2004, 14:52:34 ; Search time 1125 Seconds
 (without alignments)
 13666.591 Million cell updates/sec

Title: US-09-942-025-12
 Perfect score: 27705
 Sequence: 1 gtggcgccgatggacagaat.....cagaggatctccatgaatag 27705

Scoring table: OLIGO_NUC
 Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 200

Total number of hits satisfying chosen parameters: 3

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : Issued_Patents_NA:*
 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
 5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
 6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	14165	51.1	33529	3	US-09-144-085-3	Sequence 3, Appli
2	368	1.3	750	3	US-09-010-809-12	Sequence 12, Appl
3	220	0.8	787	3	US-09-010-809-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
 US-09-144-085-3
 ; Sequence 3, Application US/09144085

OM nucleic - nucleic search, using sw model

Run on: June 12, 2004, 14:57:19 ; Search time 6879 Seconds
(without alignments)
18383.440 Million cell updates/sec

Title: US-09-942-025-12
Perfect score: 27705
Sequence: 1 gtggcgccgatggacagaat.....cagaggatctccatgaatag 27705

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2998549 seqs, 2282253817 residues

Word size : 200

Total number of hits satisfying chosen parameters: 4

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	27705	100.0	27705	13	US-09-942-025-12	Sequence 12, Appl
2	27705	100.0	67311	13	US-09-942-025-1	Sequence 1, Appli
3	809	2.9	20922	13	US-09-942-025-14	Sequence 14, Appl
4	531	1.9	11358	13	US-09-942-025-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1

US-09-942-025-12
 ; Sequence 12, Application US/09942025
 ; Publication No. US20030054547A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kosan Biosciences, Inc.
 ; APPLICANT: Julien, Bryan
 ; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENE FROM SORANGIUM
 ; TITLE OF INVENTION: CELLULOSUM
 ; FILE REFERENCE: 30062-20020.21
 ; CURRENT APPLICATION NUMBER: US/09/942,025
 ; CURRENT FILING DATE: 2001-08-28
 ; PRIOR APPLICATION NUMBER: US 60/271,245
 ; PRIOR FILING DATE: 2001-02-15
 ; PRIOR APPLICATION NUMBER: US 09/144,085
 ; PRIOR FILING DATE: 1998-08-31
 ; PRIOR APPLICATION NUMBER: US 09/010,809
 ; PRIOR FILING DATE: 1998-01-22
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 12
 ; LENGTH: 27705
 ; TYPE: DNA
 ; ORGANISM: Sorangium cellulosum
 US-09-942-025-12

Query Match 100.0%; Score 27705; DB 13; Length 27705;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 27705; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGGCGCCGATGGACAGAATGCTTGGGTTTCTGCGCGATGCATTGGCGGATCTGCTGCAG 60
 |||
 Db 1 GTGGCGCCGATGGACAGAATGCTTGGGTTTCTGCGCGATGCATTGGCGGATCTGCTGCAG 60
 QY 61 GTTGCGCCGGGAGCCGTTTCAGGATGATGTTTCGATTCCACGAGCAGGGCCTCGACTCCGCG 120
 |||
 Db 61 GTTGCGCCGGGAGCCGTTTCAGGATGATGTTTCGATTCCACGAGCAGGGCCTCGACTCCGCG 120
 QY 121 AGGGCGCTGCTCCTGGTTGAAAAGCTGTCGAGCTGGACGGGGCGCGCTCTCCAGCCACG 180
 |||
 Db 121 AGGGCGCTGCTCCTGGTTGAAAAGCTGTCGAGCTGGACGGGGCGCGCTCTCCAGCCACG 180
 QY 181 CTGGTATGGCAGCACAGCACCCCTGAACGCGCTTGCTAGGCATCTGGCGGATACAAGCGGG 240